Research Progress in Mechanical and Manufacturing Engineering Vol. 2 No. 2 (2021) 742-760

RPMME

Homepage: http://publisher.uthm.edu.my/periodicals/index.php/rpmme
e-ISSN : 2773-4765

A Review on Assessment Quality of Fruits using Non-destructive Method

Kok Win Leong[1], **Ong Pauline**[1,*]

[1]Faculty of Mechanical and Manufacturing Engineering,
Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Johor, MALAYSIA

*Corresponding Author Designation

DOI: https://doi.org/10.30880/rpmme.2021.02.02.081
Received 30 July2021; Accepted 30 Nov.2021; Available online 25 December 2021

Abstract: Ensuring the quality of fruits will increase its productivity and marketability. However, predicting the fruit quality from human experiences is not reliable and cannot guarantee the quality of fruits because the fruit properties change rapidly during the fruit ripening process and the fruit development. Many instruments can be used to measure the total soluble solid of fruits, but mostly are destructive methods. In this paper, several rapid and non-destructive methods to determine the quality of fruits were reviewed. The result of each type of method and the obtained accuracy were presented.

Keywords: Fruit Quality, Non-Destructive Method, Total Soluble Solid

1. Introduction

In Malaysia, there are many types of tropical fruits, such as durian, watermelon, rambutans, mangosteen, langsat, pomelos and others can be easily obtained at the roadside stalls, supermarkets or hypermarket in the towns and cities when they are in season. Tropical fruits are significant to Malaysia as they are the source of income for farmers, source of vitamins, minerals and dietary fibres for humans and can generate revenue from the trade activities. Besides, after the harvesting of oil palm, paddy and rubber, the tropical fruits are usually planted in the empty land for land rehabilitation. To understand the local demand of fruits, Malaysia government will conduct yearly surveys for analysis. The indicator to indicate whether the amount of fruits can meet the needs of the people is called self-sufficiency ratio (SSR). If the SSR index of a particular fruit is more than 100%, it means that the amount of fruit harvested in a year can fully supply the need of Malaysian. From the data shown in Table 1, the supply of 11 kinds of fruits can meet the domestic demand with SSR of more than 100%. Watermelon topped the list with the highest SSR (161.3%), followed by papaya (153.1%), starfruit (132.8%), langsat (112%), jackfruit (110.5%), sweet corn (106.3%), mangosteen (105.8%), pineapple (105.2%) and others [1].

*Corresponding author: ongp@uthm.edu.my

Table 1: The Self Sufficiency Ratio (SSR) of Fruits in Malaysia year 2019 [1]

No	Types of fruits	Self-Sufficiency Ratio, (SSR) (%)
1	Watermelon	161.3
2	Papaya	153.1
3	Starfruit	132.8
4	Langsat	112.8
5	Jackfruit	110.5
6	Sweet corn	106.3
7	Mangosteen	105.8
8	Pineapple	105.2
9	Durian	105.2
10	Rambutan	101.2
11	Guava	100.0
12	Banana	98.7
13	Coconut	68.2
14	Mango	32.1

Nowadays, to provide consumers with more high-quality fruits, the quality of fruits was carefully examined. From the previous research, the quality of fruit is tested using chemicals. This process is time-consuming, laborious, and the number of testing samples at one time is very small. Therefore, a fast and accurate way to detect the quality of fruit is in need, such that a large amount of fruits can be processed at the same time for cost and time saving purposes [2]. Many physical and chemical properties can be used to determine whether the fruit has reached the mature stage, such as total soluble solid (TSS), texture starch, size, shape and others [3]. A common destructive tool to determine the sugar level content of fruit is using the refractometer. This is an optical instrument used to measure the refractive index from a few drops of fruit juice [4]. Despite the sugar content level of fruit can be easily acquired using the refractometer, it is a destructive method. Hence, in this study, some non-destructive methods to examine the quality of fruits will be discussed.

2 Non-destructive Methods to Evaluate the Quality of Fruits

The term non-destructive is defined as not causing any damage or destruction when an object or material is being tested for its properties. The non-destructive methods used to determine the quality of fruits available in the literature are discussed in the following section.

2.1 Fluorescence Method

Beaudry *et al.* [5] applied fluorescence method to determine the quality of fruits, specifically, the firmness, colour, aroma and texture. A red light with strong strength was used as a source for irradiation of fruits. The red light is shooting on the skin of the fruit and enables chlorophyll to produce a fluorescence light with higher frequency than the red light on the skin of the fruit. After that, a fluorescence detector is used to detect the fluorescence intensity produced by the skin of the fruit. Two ranges are used to classify the quality of the fruits. If the fruit fluorescence detection level is located at the integral value of first range, then it can be assumed it is a good quality of fruits and opposite in other hand. The fluorescence intensity ratio is determined by [5]:

$$\text{Fluorescense Intensity Ratio} = \frac{Fm - Fo}{Fo} \quad (1)$$

where

F_m is the fluorescence light produced by the skin

F_o is fluorescence intensity applied on the fruit skin

2.2 Computer Vision System

The computer vision (CV) system has been widely used to determine the internal quality of fruits such as determine the grade of tomatoes [6], size and weight estimation of yellow melon [7], volume and mass estimation for cherry tomato [8] and rapid determination of oil palm fresh fruit bunches maturity [9].

Generally, as shown in Figure 1, to inspect the internal quality of fruits, the utilized CV system is consisting of two parts, namely, image acquisition and image analysis. For the image acquisition, the basic components needed for a CV system are an illumination device, camera, frame-grabber, computer and high-resolution colour monitor [10]. A solid stated charged couple device (CCD) will act as the image sensor and capture the image when an object to be tested is detected. However, due to the trend of advanced technologies nowadays, a digital camera has been adopted in capturing the images of the object. By utilizing the digital camera, the time needed to convert the images to a readable format by computer can be eliminated.

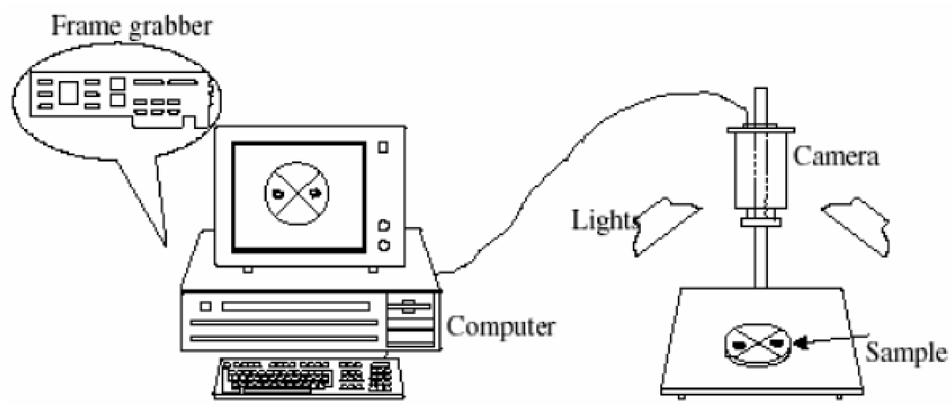

Figure 1: The component of a computer vision system [10]

For the image analysis, the raw image acquired by the digital camera will be first pre-processed. The commonly used image processing methods are image enhancement, de-noising, segmentation and feature extraction. Subsequently, further image analysis can be carried out to obtain the desired output [10].

Arakeri and Lakshmana [6] applied the CV system to determine the grade of tomatoes as ripe or unripe, and defective or non-defective. This system has two sections, fruit handling and image processing module, as shown in Figure 2.

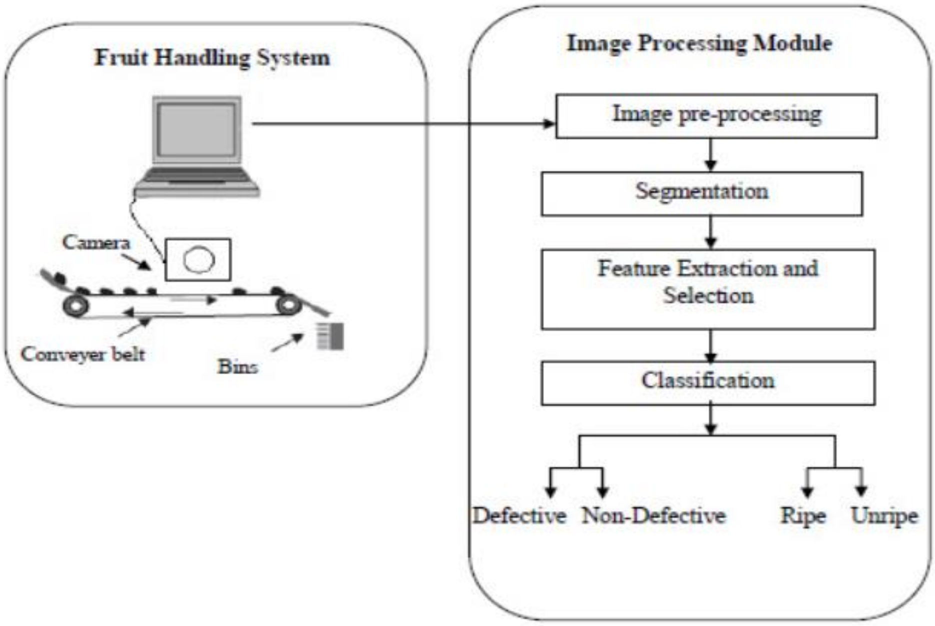

Figure 2: Tomatoes fruit grading system [6]

The fruit handling system moves the tomato from left to right using a conveyor belt with direct current gear motor as power supply. A digital camera is fixed at the centre of the conveyor belt to capture the image of the tomatoes, and followed by the image processing module. For the image processing, the median filter is used to reduce the noise and reflection on the image, so that the quality of the image can be improved. The Otsu's method is applied subsequently for image segmentation, in which the images are converted into binary images and followed by the extraction of the exact tomato region from the binary images. For feature extraction, the colour statistical and colour texture features for individual Green (G), Red (R) and Blue (B) channels are extracted to determine whether the tomatoes are defective or non-defective. Since the ripeness of the tomato is related to its colour, the calculation on mean value of R. If the mean value is more than 1, the tomato is considered as ripe and unripe for vice versa. The multilayer neural network is subsequently used to determine whether the tomato is ripe/unripe and defective/non-defective. The accuracies of 100% and 96.47% are reported for the identification of defective/non-defective and ripe/unripe classification.

Calixto et al. [7] predicted the weight and shape ratio of yellow melon through CV techniques. From Pearson's correlation coefficient model and calibration factor model, the coefficient of determination up to 0.97 and the average absolute error of 0.05kg were obtained. They also showed a success rate of 96% in shape ratio, SR-based classification by utilizing the data analysis from CV systems. Another work was conducted by Nyalala et al. [8] for the cherry tomatoes volume and mass estimation using the CV system, with the four main objectives of developing an efficient image processing algorithm, developing 2D and 3D feature extraction algorithm, creating the regression model for mass and volume estimation and investigating the relationship between tomato volume and mass. The result showed that the relationship between the volume and mass of cherry tomatoes can estimate the least accuracy of 0.9226 and the highest accuracy of 0.9706 by the developed model.

A combination of laser-light backscattering imaging and CV system experiment was done by Mohd Ali et al. [9] to evaluate the maturity of bunches oil palm fresh fruit. The RGB images were captured as the input for the CV system and the parameters obtained were further calculated using analysis of variance (ANOVA). From the result, the coefficient of determination (R^2) is more than 0.80 for the oil content and colour change of oil palm FFB. This proved that the combination of CV and laser-light backscattering imaging system is significant to examine the quality changes of oil palm in relation to the maturity [9].

2.3 Artificial Neural Network (ANN)

The ANN uses the mathematical model to imitate the working of the brain. In general, an ANN model can consist of several hidden layers. The classical artificial neuron model, as shown in Figure 3.

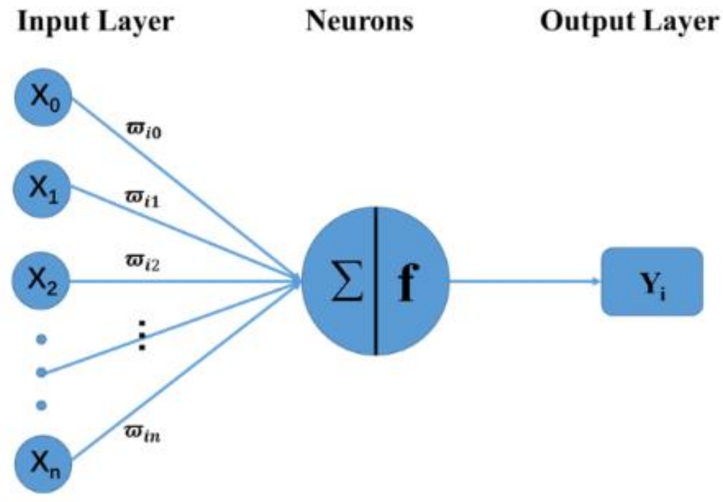

Figure 3: The classic artificial neuron model [11]

Figure 3 presents the single neuron classic artificial neuron model. Multiple neurons can exist in the hidden layer. In its simplest form, each neuron possesses a specific output function called activation function. The link between the input layers and neurons both represent a weighted value acting as the signal passing through the connection. The variation of the output network layer is dependent on how the network is built, together with weight values and activation functions [11].

Suhardiyanto et al. [12] studied the relationship between the TSS and fruit fresh weight with the electrical conductivity (EC) value of nutrient solution on every single generative stage, plant spacing and light intensity using the ANN. The generative growth of tomato fruits can be classified into three periods: flowering, fruiting and harvesting. The two different EC values with low value range (1.2-2.4 mS/cm), high value range (8.0-10.2 mS/cm) and two different planting densities of 25cm, 10cm plant spacing were considered. An ANN model with one hidden layer is used, with the input parameters are the EC values for all three generative periods, plant spacing and light intensity, while the output values are the TSS value and fruit fresh weight value. The obtained results showed that the ANN model is suitable to be used for predicting the TSS and fruit fresh weight of the tomato fruit.

Angela Vacaro de Souza *et al*. [13] used ANN modelling with variation of climax to predict the time of production of banana. In this study, six input, namely, mean, maximum and minimum temperature, precipitation, relative humidity and photoperiod were taken into account, while merely one output, *i.e*, the day to harvest, was considered. The data obtained from the ANN was compared with the desired data and it is found that both values were very similar. The training ANN also performed well with only 0.00397 for the error and attained up to 0.8998 for the coefficient of determination, proving the reliability of this model.

2.4 Near Infrared Reflected Spectroscopy (NIRS)

In the past research, NIRS has been used to measure the quality of fruits, with a reasonably good accuracy. Many properties can be measured by applying the NIRS, including the TSS, carbohydrate, vitamin C and other related chemical properties. The concept of NIRS instruments is based on the electromagnetic radiation interaction principle and natural products. Basically, the light source used to release the electromagnetic radiation waves and wavelength produced will be examined.

Rungpichayapichet *et al*. [14] applied the NIRS to predict the ripeness of the mango fruit in the postharvest stage. A portable vis/NIR photo-diode array spectrometer equipped with a MMS1 module addition with a fibre optic interactance probe was applied to obtain NIRS spectra in reflectance mode. The collected spectra data were within 700 to 1100nm at 2 nm intervals with resolution of 10 nm. The partial least square (PLS) regression was used to establish the calibration model based on the spectra range between 700-1100nm. The obtained results showed that the NIR spectra has a strong correlation with the TSS value of mango fruits, with the values $R^2 > 0.8$ and SER < 1.3% were observed.

Ebrahiema Arendse *et al.* [15] proposed fourier transform near infrared (FT-NIR) diffuse reflectance spectroscopy as a non-destructive method to analyse the external and internal quality of intact pomegranate fruit. Through this method, there are two different ways for spectral acquisition, which is using a rotating in integrating sphere (IS) and contact with the sample directly, or utilizing an optical fibre couple emission head (EH) for contactless measurement. The spectra data obtained from NIR will be analysed using the principle component analysis (PCA) and PLS. From the result, it was shown that the EH spectra provided better prediction results for TSS, pH and vitamin C, while IS spectra is suitable for determining the colour component of pomegranate fruits. The calibration models by using FT-NIR diffuse reflectance spectroscopy successfully predicted internal and external quality of pomegranate fruits due to precise examination of the parameter that existed on the fruit.

2.5 Hyperspectral Imaging (HIS)

HSI measures the wavelength groups of an organics product. The HSI information is built either for 3-dimensional (3D) or 2-dimensional (2D) examination in the image. For 2D, the axis is x and y – axis where 3D is x, y and λ for the wavelength axis. Consequently, a hyperspectral image is made out of 100 coterminous wavebands for each spatial area focused on the object. Every pixel in the HSI helps to hold the spectrum of the specific position. Some data can be obtained through a particular pixel after the spectrum is acquired. HSI can be seen as a combination of spectroscopy and imaging or computer vision. Along these lines, HSI not exclusively can decide the chemical and properties covered in an organics product but also for the size, appearance, and shading.

Yan-Ru Zhao *et al.* [16] studied the application of HSI for assessment of TSS in mulberries. A HIS apparatus (Figure 2.4), consisting of illumination source, a mobile platform and a computer with the spectral imaging system software, was used to acquire the hyperspectral images of mulberry samples. A spectral data matrix of the dimension 310×460 (sample × wavebands) was obtained, and transferred to the least square support vector machine (LS-SVM) regression model and PLS regression model for further analysis. From the obtained results, these two models were successful in prediction of TSS of mulberry.

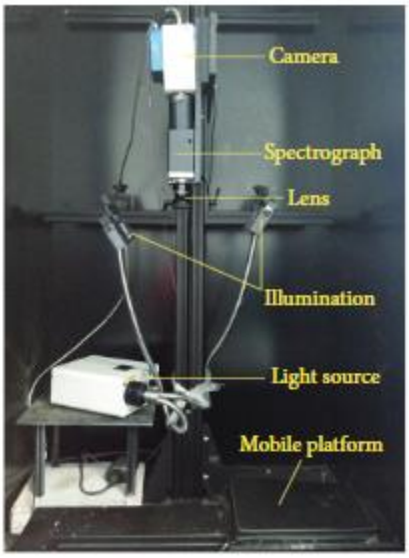

Figure 4: Apparatus in HIS [16]

Pullanagari and Li [17] assessed the firmness and TSS of sweet cherries using the HIS. The image acquisition used to capture the sample image is by using a push-broom HIS camera. The NIR spectra is within the range of 500 – 1600 nm and the spatial information is about 320 × 500 pixels. The quality parameters that are concerned in this experiment is TSS and flesh firmness. The data from the image acquisition were further used to evaluate the prediction performance PLS regression and Gaussian process regression (GPR). The results showed that the GPR model gave higher accuracy with a prediction interval coverage probability (PICP) of 0.90 – 0.97.

2.6 Laser Light Backscattering Imaging (LLBS)

LLBI can be utilized to quantify TSS, shading, firmness, titratable acidity, and dampness substance of organic products. The properties of organic products can be depicted by utilizing the profile created on light absorption and light-scattering dissipation. The chemical substances such as moisture, soluble solids, colour content and others are required during the light absorption process. However, the scattering of organic products not merely depends on chemical properties but some of the physical properties such as density or size of cell also can be determined. For the fruits, the backscattering appearance on the binding surface of the cell can be found easily and ordinary because of movable particles which always influence the reflecting coefficient. But some immobile particles like starch or chloroplasts perhaps to happen scattering by refracting occur around their surfaces. The remaining light either will be absorbed by the fruit tissues or emitted out the fruit. Amount of light can be absorbed is differ to each organic product and truly dependent on some properties of the product or wavelength of light. Hence, organic tissues might be considered on the effect of light absorption and scattering on each specific organic product [18].

Qing *et al.* [19] used LLBI to predict the TSS of apples. Results showed that the $R^2 = 0.656$ and SEP = 0.125 were obtained based on the PLS regression models. In their study, three different segmentation methods were considered, namely, bimodal threshold algorithm, iterative arithmetic and the Otsu method. The Otsu method was reported to show higher accuracy in predicting the TSS.

Nurazwin Zulkifli *et al.* [20] applied the LLBI for the ripeness stage prediction for the “Berangan" pineapple fruit. The stages of pineapple ripeness from 2 to 7 were selected for backscattering image acquisition by a CCD camera together with a laser diode emitting light of 658nm wavelength. From the captured backscattering images, the field of backscattering area and grey level intensity were extracted and used for estimation parameter quality of the pineapple fruit. The result showed the coefficient of determination (R^2) of the total soluble content exceeds 0.7, indicating that there was a correlation between the LLBI parameter and the colour of pineapple, which could be used for prediction the maturity stages of pineapple fruits. In addition, the extracted data was further evaluated by statistical analysis of linear discriminant analysis (LDA). The results also showed a correct classification percentage of 94.2%, indicating the potential of the LLBI for the pineapple maturity classification purpose.

2.7 Magnetic Resonance Imaging (MRI)

Nuclear magnetic resonance (NMR) is related to the study on nuclear physics. MRI is an extraordinary imaging innovation with re-constructive signals produced by the nuclei of NMR. A magnetic field plate is prepared and organic products are arranged on the plate. Radio–frequency pulses are utilized to activate the hydrogen nuclei in the organic products. Hydrogen atoms recruited the energy in the presence of nuclear resonance created by the radio–frequency pulses. When switching off the radio–frequency pulses, a portion of hydrogen nuclei sends out a radio signal through emission of energy. Besides, the absorption of energy is used to get images processed by the computer [18].

Clark *et al.* [21] studied the relationship between TSS of kiwifruit and relaxation change during the growth and maturity through the quantitative NMR imaging. The result showed that when changing the relaxing parameters, the TSS will increase up to 125%, due to the impact of cell structure and magnetic field strength during the relaxation process.

Maja Musse *et al.* [22] used MRI to study the structural aspects of the tomato fruit. A 0.2-T electromagnet scanner was used for image acquisition method. The multiple gradient echo images were used to evaluate the air bubble content in tomato tissue. The microstructure of tomato was further investigated by measuring spin-spin and spin-lattice relation time distribution. The relationship between the air bubble content with tissues was determined from the MRI image, and was found able to distinguish various types of tissues in the tomato fruit.

2.8 Convolutional Neural Network (CNN)

Deep Learning (DL) is an advanced technique used for large quantities of image processing in a short time. CNN is one of the subfields in the DL. Compared to the aforementioned ANN, CNN is considered as the improved version of ANN and has been widely used in computer vision applications. CNN is able to solve more complex problems in a short time due to a more complicated model has been constructed. In contrast, to solve the same problem, large time consuming is needed for ANN. The architecture of CNN consists of several layers, including the convolutional layers, acting as feature extractor from input image, pooling layers, working on reducing the dimension of image, fully connected layer, acting as a classifier and the last layer acts as a higher level analyser on the extracted feature and provides detail information as an output [23]. An example of CNN architecture is shown in Figure 5.

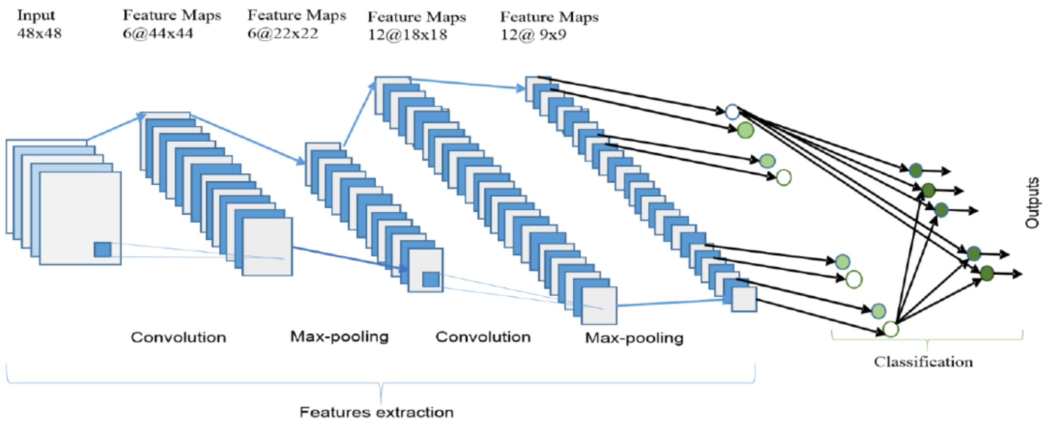

Figure 5: The overall architecture of the CNN [24]

In terms of application of CNN in agriculture, Vijayalakshmi *et al.* [25] used CNN to classify the banana species, in which the accuracy of 96.98% was reported. The obtained results outperformed the classifiers of K-nearest neighbour and Histogram of Gradient model.

Agarwal *et al.* [26] performed the experiment of tomato leaf disease detection using the application of CNN. Ten thousand images of tomato disease leaf were used for training dataset, 7000 images for validation data set and 500 images for testing dataset. After that, a CNN model was established with 3 convolution and 3 max pooling layers, followed by 2 fully connected layers, as shown in Figure 6. The results showed that the accuracy model was in the range of 76% and 100% for each class. The average testing accuracy of the model was 91.2% for 9 diseases and l health class.

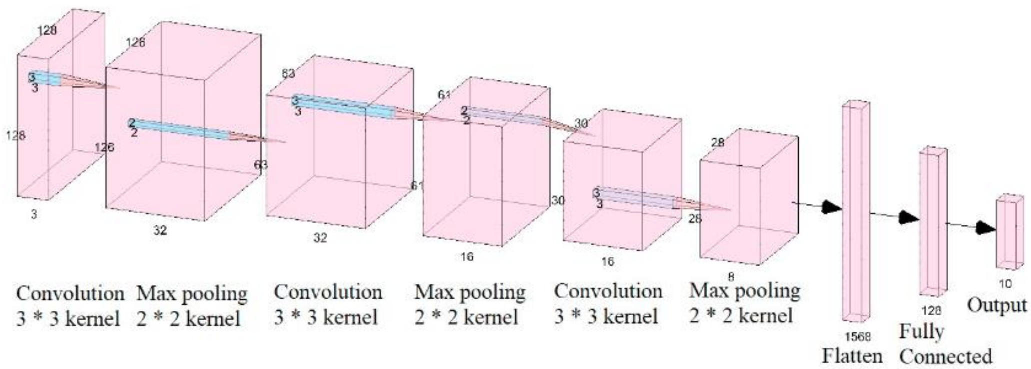

Figure 6: The proposed convolution network

Cecotti *et al.* [27] utilized the CNN for grape detection, in which two types of grapes were considered. Several types of images, namely, colour images, grayscale images, and colour histogram were used to evaluate the impact of the input feature space for classification performance of CNN. The results showed that the segmentation of grape images can be done more effectively using feature spaces and colour images. The CNN was then transferred to Resnet networks, reaching an accuracy of 99% for both types of grapes.

2.9 Support Vector Machine (SVM)

SVM is a mathematical model established to solve classification and regression problems [28]. SVM can act as classical machine learning techniques for solving the huge data classification problems and widely used in multi domain applications [29]. The SVM is based on the concepts of mapping the vectors into some high dimensional feature space through some non-linear mapping chosen as preferred [30]. A straight line decision surface is built towards the space with unique properties to confirm high generalization ability of the network. Figure 7 shows an example of a separable problem in a 2D space using the SVM.

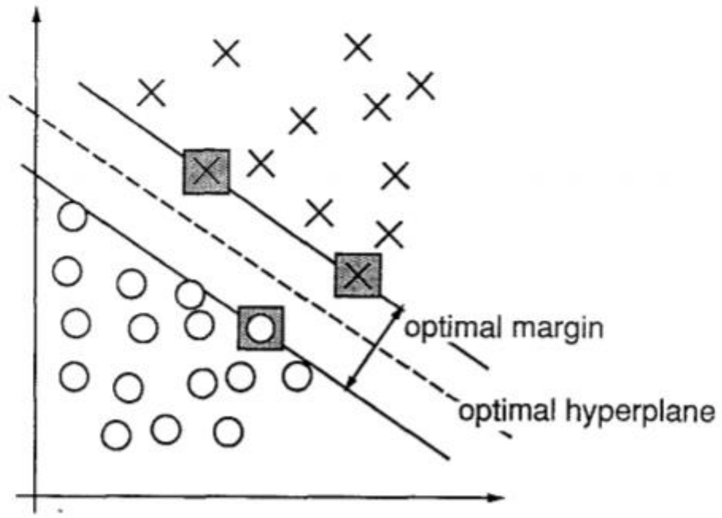

Figure 7: An example of a separate problem in a 2D space using a SVM. The grey squares defines as a support vector [30]

In order to make a decision surface corresponding to a two degree polynomial, it can create a feature space Z, consists $N = n(n+3)/2$ coordinates of the form;

$z_1 = x_1, ..., z_n = x_n$ n coordinates,

$z_{n+1} = x_1^2, ..., z_{2n} = x_n^2$ n coordinates,

$z_{2n+1} = x_1 x_2, ..., z_N = x_n x_{n-1}$ $\frac{n(n-1)}{2}$ coordinates,

where $x = (x_1, ..., x_n)$. The hyperplane will be constructed in this space.

Sun *et al.* [31] studied the relationship between fruit moving speed and the TSS of 'Cuiguan' pears, The results for R^2 and RMSEP of 0.850 and 0.557 were reported for the LS-SVM model, while 0.916 and 0.530 were obtained for the PLS regression model using the original spectra data. Their experiments showed that the fruit moving speed within the range of 0.3 – 0.7 ms^{-1} will slightly affect the performance of the model. The speed of 0.5 ms^{-1} is reported as the best model to determine the TSS is PLS.

Azarmdel *et al.* [32] studied the ripeness level of the mulberry using ANN and SVM. Initially, the mulberry fruit's images were captured by an imaging system. The quality parameters of mulberry fruit, specifically, the geometrical properties, colour and texture characteristics, were then extracted using correlation-based feature selection (CFS) and consistency subset (CONS) for feature reduction purposes. CFS and CONS set data underwent analysis of ANN model and SVM model, respectively.

For the SVM, six groups of fruit were classified based on two different feature reduction methods by the SVM algorithm such as MSE values, classification values, RBF, polynomial and linear kernel function. Moreover, the optimal value of the penalty parameter (C) and the kernel parameter (σ) also has been determined by calculating the MSE values together with classification accuracy. The result shows that MSE values of the CFS and CONS subset have the lowest value of $9 \times^{-3}$ for both training and test sets, proving the reliability of the SVM classifier to classify the maturity level of mulberry based on the colour, shape and texture of fruits.

2.10 Random Forest (RF)

RF is a constructed model that operates rapidly in the mathematical algorithm. The accumulated ideas are used to build the RF model for both regression and classification problems. The principle of RF is based on phenomenon growing of trees in a random perturbation situation. It tries to sort out the problems by using numerous models fitted to the same dataset [33]. RF work as gathering a set of random trees, construct over n_{tree} bootstraps sample $L^1 \ldots L^{ntree}$ of the training set L [34]. The trees of RF able to rapid computations and preserve statistical performance quickly. The basic process works on the RF which is; a fixed number of variables is picked randomly to find the most suitable split in every single node, all the trees of the forest are maximal trees due to not be pruned . The resulting pattern is accumulating all the estimators resulting from any relevant trees, instructed by $\widehat{f_1}, \ldots, \widehat{f_{n_{tree}}}$. From that, the accumulation also functions as calculating the average prediction value and making a prediction at new point of x.

$$\hat{f}_{RF}(x) = \frac{1}{n_{tree}} \sum_{k=1}^{n_{tree}} \hat{f}_k(x) \tag{2}$$

Santos Pereira *et al.* [35] predicted the maturity of papaya fruit with digital imaging and RF. The images were acquired from the CV system, consisting of three components: an illumination source, a consumer digital camera and a rectangular chamber with matte back internal walls to prevent specular reflections. After that, two datasets were created from the extracted data, namely, the cross-validation and prediction test set. The cross-validation set was used for adjusting the hyper-parameters of the model, while the prediction set was used to examine the classification performance. From the result, the cross-validation set achieved performance of classification with a percentage of 94.3%, where the prediction set showed 94.7% for calibrating the classification correctly.

Fukuda *et al.* [36] used RF to relate the peel and colour to the quality of fresh mango fruits. For the sample, the mango was collected from two mango cultivars with different maturity stages. To conduct the experiment, three different temperature treatments were prepared to alter the postharvest fruit quality and its peel colour. The peel colour was measured based on the CIE $L^* a^* b^*$ colour system. Six RF models were established to evaluate the hardness, ascorbic acid and TSS of the mango fruits. Each of the models went through the modelling process 10 times repeatedly using ten different sets of initial conditions to investigate the variability of the model structures by examining the output data of the model. From the result, it can be said that RF models were able to connect the peel colour parameter to every single fruit quality when reached maturity. This is because the variability in model structure was so small even though based on the result from 10 different sets of initial conditions. The standard deviation of the performance measures also showed a very small value of error, which proved the reliability of the RF model.

De Miranda Ramos Soares *et al.* [37] applied the RF to predict the basic-dye biosorption process for the orange fruits. Three different RF models were established separately to determine the prediction value of final concentration of methylene blue (C_f in mg.L^{-1}), adsorption capacity (Q in mg.g^{-1}) and the adsorbate removal percentage (R %). For the input variables for these models, several parameters such as Temperature (^{0}C), pH, adsorbent dosage (% w.w^{-1}), contact time (min), salinity (g.L^{-1} of NaCl), initial methylene blue concentration (Ci mg.L^{-1}) and rotation (rpm) were determined from the numerous orange fruit samples. Figure 8 shows the schematic diagram of the three models developed using RF prediction. In order to validate the RF models, the values of coefficient of determination (R^2) and mean square error (MSE) were calculated. The RF models are also compared with convolutional ANN using the similar way to build the ANN models. From the result, it seems that the RF and ANN models exhibited similar performances as the values of R^2 were 0.9739 and 0.9734, 0.9932 and 0.9919 for C_f, 0.9318 and 0.9257 for R % obtained from both models.

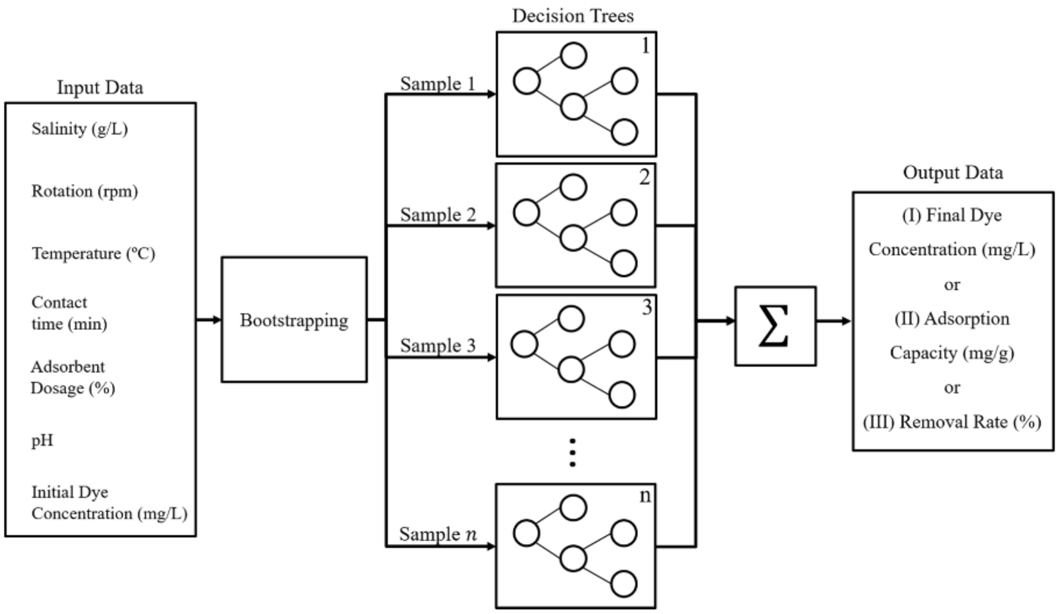

Figure 8: The schematic diagram of the three models developed using RF prediction [37]

Lastly, Table 2 summarizes the non-destructive methods used to determine the quality of fruits in previous studies.

Table 2: Summary of previous study on non-destructive method in fruits quality assessment system

No	Non-destructive method	Agriculture area	Sample used	Precision	Classifier	Framework	Compare with other method	Ref.
1	Computer Vision System	Tomato fruit grading system	520 images captured by digital camera with size 512×512 pixels	100% (CA) for defective/non-defective 96.47% (CA) for ripe/unripe	ANN neural network	Standard (three layer feed forward neural network)	Statistical features extracted from the tomato images was compared with the texture features using R, G, and B values of the tomato images.	[6]
		Development size and weight estimation of yellow melon model	105 samples of melon images captured by digital camera with 4608×3072 pixels. Digital weight machine used to determine the weight of melon in kilograms	96% (CA)	Person's correlation of melon classified into four groups according to their computer vision real shape ratio (SR_{real})	Developed by authors	N/A	[7]
2	Artificial Neural Network	Prediction of bananas' harvest time with variables on climatic	Climatological data were measured by automatic stations	89% (CA)	ANN	Multilayer Perceptron (MLP)	N//A	[13]
3	Near Infrared Reflected Spectroscopy	Quality prediction for the ripeness of mango fruits in the postharvest stage	Fruits spectrum were measured using NIR photo-diode array spectrometer	80%(CA)	Ripening index (RPL) model	Partial least square (PLS) regression analysis	N/A	[14]

4	Hyperspectral Imaging	Quality assessment of TSS in mulberries	Hyperspectral images of 310 mulberries were acquired by CCD camera	89% (CA) 95.6%(CA)	RF-PLSR RF-LS-SVM	Developed by authors	LS-SVM regression models which is nonlinear regression model had better performance for predicting TSS of mulberries	[16]
5	Convolutional Neural Network	Types of grape fruits detection	Two grape types images were captured using built-in camera from iPhone 8	99%(CA)	CNN	Resnet networks	N/A	[27]
6	Support Vector Machine	Determine the ripeness level of mulberry	Six different classes of ripeness mulberry images were captured by CCD camera. Images will further undergo feature extraction using two feature reduction methods, Correlation-based Feature Selection subset (CFS) and Consistency subset (CONS)	CFS subset ANN 100% (CA) 100% (CA) 99.1% (CA) CONS subset ANN 100% (CA) 98.9% (CA) 98.3% (CA) CFS subset SVM 99.78% (CA) 99.12% (CA) CONS subset SVM 99.57% (CA) 98.25% (CA)	ANN SVM	<u>ANN</u> 20 neurons in the hidden layer 15 neurons in hidden layer <u>SVM</u> Three different kernel functions (RBF, polynomial, and linear)	ANN showed a significant advantage over the SVM for the mulberry classification	[32]

7	Random Forest	Prediction for maturity of papaya fruit	57 fruits image were acquired using digital camera. Images will further undergo feature extraction from each colour channel with a total seven group of data.	94.3% (CA) 94.7% (CA)	Random Forest	RF with two datasets (cross-validation and prediction set)	N/A	[35]

** CA = Classification Accuracy

3. Conclusion

There is a high demand in market and industrial sectors to produce better quality of fruits. Traditionally, the fruit growers use their experience to estimate the right time for harvesting based on the color changes of the fruits. Unfortunately, this method is not reliable and cannot ensured the quality of fruit because the properties of fruits itself change rapidly during the fruit ripeness process and fruit development. Moreover, there were existed many issues affecting the maturity of the fruit, such as climate variation, air temperature, soil moisture, soil fertilizer, air relation and humidity. Hence, the development of a rapid and non-destructive method to predict the internal quality of fruits will give a better estimation method and the wastage of fruits also can be minimized. By using the non-destructive method, there is no need to utilize the special equipment and chemical analysis to estimate the TSS of fruits. Hence, the evaluation time of the inner quality of fruits can be reduced, in addition to a cheaper processing cost could be obtained. Lastly, rapid and non-destructive method ensures the competitiveness in the market and will be widely used in the future.

References

[1] Department of Statictics Malaysia. *Supply and Utiliation Accounts Selected Agricultural Commodities, Malaysia*. 2015-2019; Available from: https://www.dosm.gov.my/v1/index.php?r=column/cthemeByCat&cat=164&bul_id=OTM1TDMzS1IvYm5mU1JiU1Fwekt3UT09&menu_id=Z0VTZGU1UHBUT1VJMFlpaXRRR0xpdz09.

[2] Alia, M.M., Hashima, N., Abd Aziza, S., Lasekanb, O.J.J.o.A., & Engineering, F. (2020). An overview of non-destructive approaches for quality determination in pineapples. 1, 0011.

[3] Li, B., Lecourt, J., & Bishop, G.J.P. (2018). Advances in non-destructive early assessment of fruit ripeness towards defining optimal time of harvest and yield prediction—a review. 7, 3.

[4] Villaseñor-Aguilar, M.-J., Sánchez-Bravo, M.-G., Padilla-Medina, J.-A., Vázquez-Vera, J.L., Guevara-González, R.-G., García-Rodríguez, F.-J., & Barranco-Gutiérrez, A.-I.J.A.S. (2020). A Maturity Estimation of Bell Pepper (Capsicum annuum L.) by Artificial Vision System for Quality Control. 10, 5097.

[5] Beaudry, R.M., Armstrong, P.R., Song, J., & Deng, W., *Non-destructive method and apparatus for detection of fruit and vegetable quality*. 1998, Google Patents.

[6] Arakeri, M.P. & Lakshmana (2016). Computer Vision Based Fruit Grading System for Quality Evaluation of Tomato in Agriculture industry. Procedia Computer Science, 79, 426-433.

[7] Calixto, R.R., Pinheiro Neto, L.G., Cavalcante, T.d.S., Aragão, M.F., & Silva, E.d.O. (2019). A computer vision model development for size and weight estimation of yellow melon in the Brazilian northeast. Scientia Horticulturae, 256, 108521.

[8] Nyalala, I., Okinda, C., Nyalala, L., Makange, N., Chao, Q., Chao, L., Yousaf, K., & Chen, K. (2019). Tomato volume and mass estimation using computer vision and machine learning algorithms: Cherry tomato model. Journal of Food Engineering, 263, 288-298.

[9] Mohd Ali, M., Hashim, N., & Abdul Hamid, A.S. (2020). Combination of laser-light backscattering imaging and computer vision for rapid determination of oil palm fresh fruit bunches maturity. Computers and Electronics in Agriculture, 169, 105235.

[10] Narendra, V. & Hareesha, K.J.I.J.o.C.A. (2010). Prospects of computer vision automated grading and sorting systems in agricultural and food products for quality evaluation. 1, 1-9.

[11] Feng, C., Zhao, N., Yin, G., Gan, T., Yang, R., Chen, X., Chen, M., & Duan, J. (2021). Artificial neural networks combined multi-wavelength transmission spectrum feature extraction for sensitive identification of waterborne bacteria. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 119423.

[12] Suhardiyanto, H., Arif, C., Setiawan, B.I.J.J.o.M., & Sciences, F. (2009). Optimization of EC values of nutrient solution for tomato fruits quality in hydroponics system using artificial neural network and genetic algorithms. 41, 38-49.

[13] de Souza, A.V., Bonini Neto, A., Cabrera Piazentin, J., Dainese Junior, B.J., Perin Gomes, E., dos Santos Batista Bonini, C., & Ferrari Putti, F. (2019). Artificial neural network modelling in the prediction of bananas' harvest. Scientia Horticulturae, 257, 108724.

[14] Rungpichayapichet, P., Mahayothee, B., Nagle, M., Khuwijitjaru, P., & Müller, J. (2016). Robust NIRS models for non-destructive prediction of postharvest fruit ripeness and quality in mango. Postharvest Biology and Technology, 111, 31-40.

[15] Arendse, E., Fawole, O.A., Magwaza, L.S., Nieuwoudt, H., & Opara, U.L. (2018). Fourier transform near infrared diffuse reflectance spectroscopy and two spectral acquisition modes for evaluation of external and internal quality of intact pomegranate fruit. Postharvest Biology and Technology, 138, 91-98.

[16] Zhao, Y.-R., Yu, K.-Q., & He, Y.J.J.o.a.m.i.c. (2015). Hyperspectral imaging coupled with random frog and calibration models for assessment of total soluble solids in mulberries. 2015,

[17] Pullanagari, R.R. & Li, M. (2021). Uncertainty assessment for firmness and total soluble solids of sweet cherries using hyperspectral imaging and multivariate statistics. Journal of Food Engineering, 289, 110177.

[18] Li, J.L., Sun, D.W., Cheng, J.H.J.C.R.i.F.S., & Safety, F. (2016). Recent advances in nondestructive analytical techniques for determining the total soluble solids in fruits: a review. 15, 897-911.

[19] Qing, Z., Ji, B., & Zude, M. (2007). Predicting soluble solid content and firmness in apple fruit by means of laser light backscattering image analysis. Journal of Food Engineering, 82, 58-67.

[20] Zulkifli, N., Hashim, N., Abdan, K., & Hanafi, M. (2019). Application of laser-induced backscattering imaging for predicting and classifying ripening stages of "Berangan" bananas. Computers and Electronics in Agriculture, 160, 100-107.

[21] Clark, C.J., Drummond, L.N., MacFall, J.S.J.J.o.t.S.o.F., & Agriculture (1998). Quantitative NMR imaging of kiwifruit (Actinidia deliciosa) during growth and ripening. 78, 349-358.

[22] Musse, M., Quellec, S., Devaux, M.-F., Cambert, M., Lahaye, M., & Mariette, F. (2009). An investigation of the structural aspects of the tomato fruit by means of quantitative nuclear magnetic resonance imaging. Magnetic Resonance Imaging, 27, 709-719.

[23] Kamilaris, A. & Prenafeta-Boldú, F.X.J.T.J.o.A.S. (2018). A review of the use of convolutional neural networks in agriculture. 156, 312-322.

[24] Alom, M.Z., Taha, T.M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M.S., Van Esesn, B.C., Awwal, A.A.S., & Asari, V.K.J.a.p.a. (2018). The history began from alexnet: A comprehensive survey on deep learning approaches.

[25] Vijayalakshmi, M. & Peter, V.J.J.I.J.o.I.T. (2020). CNN based approach for identifying banana species from fruits. 1-6.

[26] Agarwal, M., Singh, A., Arjaria, S., Sinha, A., & Gupta, S. (2020). ToLeD: Tomato Leaf Disease Detection using Convolution Neural Network. Procedia Computer Science, 167, 293-301.

[27] Cecotti, H., Rivera, A., Farhadloo, M., & Pedroza, M.A. (2020). Grape detection with convolutional neural networks. Expert Systems with Applications, 159, 113588.

[28] Noble, W.S.J.N.b. (2006). What is a support vector machine? 24, 1565-1567.

[29] Suthaharan, S., *Support vector machine*, in *Machine learning models and algorithms for big data classification*. 2016, Springer. p. 207-235.

[30] Cortes, C. & Vapnik, V.J.M.l. (1995). Support-vector networks. 20, 273-297.

[31] Sun, T., Lin, H., Xu, H., & Ying, Y. (2009). Effect of fruit moving speed on predicting soluble solids content of ‘Cuiguan’ pears (Pomaceae pyrifolia Nakai cv. Cuiguan) using PLS and LS-SVM regression. Postharvest Biology and Technology, 51, 86-90.

[32] Azarmdel, H., Jahanbakhshi, A., Mohtasebi, S.S., & Muñoz, A.R. (2020). Evaluation of image processing technique as an expert system in mulberry fruit grading based on ripeness level using artificial neural networks (ANNs) and support vector machine (SVM). Postharvest Biology and Technology, 166, 111201.

[33] Naroui Rad, M.R., koohkan, S., Fanaei, H.R., & Pahlavan Rad, M.R. (2015). Application of Artificial Neural Networks to predict the final fruit weight and random forest to select important variables in native population of melon (Cucumis melo L.). Scientia Horticulturae, 181, 108-112.

[34] Antoniadis, A., Lambert-Lacroix, S., & Poggi, J.-M. (2021). Random forests for global sensitivity analysis: A selective review. Reliability Engineering & System Safety, 206, 107312.

[35] Santos Pereira, L.F., Barbon, S., Valous, N.A., & Barbin, D.F. (2018). Predicting the ripening of papaya fruit with digital imaging and random forests. Computers and Electronics in Agriculture, 145, 76-82.

[36] Fukuda, S., Yasunaga, E., Nagle, M., Yuge, K., Sardsud, V., Spreer, W., & Müller, J. (2014). Modelling the relationship between peel colour and the quality of fresh mango fruit using Random Forests. Journal of Food Engineering, 131, 7-17.

[37] de Miranda Ramos Soares, A.P., de Oliveira Carvalho, F., de Farias Silva, C.E., da Silva Gonçalves, A.H., & de Souza Abud, A.K. (2020). Random Forest as a promising application to predict basic-dye biosorption process using orange waste. Journal of Environmental Chemical Engineering, 8, 103952.